Purdue University
Purdue e-Pubs

Purdue CIBER Working Papers

Krannert Graduate School of Management

1-1-2010

A Symmetry Hypothesis and Measurement Biases in the Factor Content of Trade

Laura Puzzello
Purdue University

PURDUE CIBER

Center for International
Businesses Education and Research

A Symmetry Hypothesis and Measurement Biases in the Factor Content of Trade

Laura Puzzello
Purdue University

CIBER Working Paper Series
2010-002

Purdue University
Krannert Building
403 W. State Street
West Lafayette, IN 47907-2056
(765) 496-6779
Fax (765) 494-9658

A Symmetry Hypothesis and Measurement Biases in the Factor Content of Trade

Laura Puzzello*†

Purdue University

Abstract

I revisit Reimer (2006), and Trefler and Zhu (2005, 2006) (RTZ) tests of the Vanek proposition in the presence of international differences in production techniques and global production sharing. In this framework, knowing the bilateral details of each country's input-output structure is key to the correct calculation of the factor content of trade. Because input-output tables typically lack this detail, RTZ impute the relevant input-output coefficients using a method that implicitly assumes that international flows of goods respond to trade determinants independently of their end-use (*Symmetry Hypothesis*). This paper uses survey-based input-output coefficients from the Asian Input-Output tables (AIO) that do provide bilateral detail. Exploiting methodological differences in the compilation of the AIO tables and the data underlying RTZ studies, I empirically test the symmetry hypothesis and find that it fails. This failure causes input-output data imputed following RTZ methodology to overstate the gross quantity of both domestic and foreign factors' services embodied in a country's trade. However, both biases are systematic and tend to cancel each other out resulting in only a small positive bias on net flows of factors and in the performance of the Vanek proposition.

Keywords: Factor content; International differences in techniques of production; Patterns of trade

JEL classification: F1; F11; F14

*Address: Purdue University, Department of Economics, West Lafayette, IN 47907; e-mail: lpuzzell@purdue.edu; phone: +1-765-743-3095; fax: +1-765-494-1526.

†I am extremely grateful to David Hummels for his many questions, invaluable comments and constant support. I thank Chong Xiang, Vova Lugovskyy, Kevin Mumford and Justin Tobias for very helpful discussions. My thanks to Sirsha Chatterjee, Anca Cristea, Tasneem Mirza, Iryna Topolyan and participants at Midwest International Trade Meetings (Iowa City, Spring 2009), Purdue and IDE-JETRO Conference on Vertical Specialization and International Trade (West Lafayette, Fall 2008), and seminars at Purdue University. I am also indebted to Yoko Uchida for help with the Asian data and Terry Walmsley for granting me access to the GTAP database. All remaining mistakes are my own.

Introduction

The Vanek (1968) proposition establishes a simple, yet elegant, linear relationship between trade in factor services and countries' factor endowments. That is, a country relatively well-endowed with labor will be a net exporter of labor services embodied in traded commodities. Early empirical studies found scarce support for the Vanek proposition in its original formulation[1] (Leontief, 1953; Maskus, 1985; Bowen et al., 1987). This motivated amendments of the standard Heckscher-Ohlin-Vanek (HOV) theory on the production and on the absorption side[2].

Recent contributions to the literature emphasize differences across countries in techniques of production which arise when factor prices are not equalized[3]. Failures of factor price equalization are especially problematic when intermediate inputs are traded internationally. In this case, proper measurement of factor services requires the researcher to distinguish the contribution of individual source countries to the production process because the same input may embody different factor services depending on its origin[4]. Reimer (2006) and Trefler and Zhu (2005, 2006) (RTZ) formalize this intuition in the correct definition of factor content of trade to be applied in HOV frameworks with traded inputs and unequal factor prices. Empirical estimates of these models reveal an improved performance of the Vanek proposition.

Most input-output tables provide data on the quantity of each input necessary to produce a final output but do not specify the source of those inputs. Since source information is critical to testing the amended Vanek proposition, RTZ employed a simple imputation procedure to distribute total input use across sources. Suppose for example that the U.S. imports \$60 ml in auto industry products from Japan and its total use of auto products is \$200 ml. The RTZ imputation procedure would then assume that each American industry relies on Japanese auto industry products for 30% of its total use of

[1] International data show that trade in factor services is largely overestimated by the theory and, while rich countries are scarce in most factors, poor countries are abundant in most factors (Trefler, 1995).

[2] See Trefler (1993, 1995), Davis and Weinstein (2001, 2003).

[3] Repetto and Ventura (1998) bring evidence that factor prices differ internationally even after controlling for factor-augmenting technology differences across countries. Romalis (2004) finds that locally abundant factors are relatively inexpensive (net of neutral technology effects).

[4] Empirical evidence shows that a country's production techniques intensively use that country's abundant factors. See Davis and Weinstein (2001), Xiang (2007) and Schott (2003).

auto parts. The same ratio is used whether auto products traded between Japan and the U.S. are for final consumption, intermediate use or some combination of the two. In other words, RTZ's imputation implicitly assumes that international flows of goods respond to trade determinants independently of their end-use (*Symmetry Hypothesis*).

I revisit RTZ calculations using survey-based input-output data for nine East Asian countries and the U.S. from the Asian Input-Output (AIO) tables that provide direct information on the sources of inputs. These data allow me to test the symmetry hypothesis in RTZ and to conduct the appropriate Vanek test using better data. In particular, I take the AIO tables as the best available approximation to each country's true input-output structures and show that the imputation methodology underlying the data used in RTZ generates input-output coefficients that deviate substantially from the AIO data. Importantly, these differences are not "noise", and can be explained by trade determinants such country size, factor endowments and trade costs. In short the symmetry hypothesis fails empirically.

To verify how failures of the symmetry hypothesis affect the performance of the Vanek proposition under failures of factor price equalization, I compare the estimates of the HOV model based on the AIO data to estimates based on the input-output coefficients imputed following RTZ methodology. In both cases I find that the empirical performance of the Vanek proposition improves when the standard model is amended to account for international differences in production techniques and trade in intermediates. Surprisingly, even though the AIO data differ substantially from the input-output coefficients imputed à la RTZ, the empirical performance of the Vanek proposition looks similar in the two cases. A decomposition analysis of the measured factor content of trade reconciles this inconsistency. I show that data imputed using the RTZ methodology inflate the gross quantity of both domestic and foreign factor services in trade. These biases are systematic and tend to cancel each other out resulting in only a small positive bias on the performance of the Vanek proposition when factor prices are not equalized.

In testing the symmetry assumption this paper relates to the fragmentation literature and, more closely, to a new line of research interested in exploring determinants of trade in intermediates and final goods (Sitchinava et al., 2009; Bergstrand and Egger, 2008).

While Bergstrand and Egger derive different theoretically-motivated gravity specifications for trade in inputs and final goods, Sitchinava et al., using detailed data on U.S. imports, show that these two types of flows respond differently to countries' factor endowments, quality of contractual environment and market thickness. My results are consistent with Sitchinava et al. preliminary findings.

Failures of the symmetry hypothesis imply that comparative advantage forces asymmetrically affect patterns of trade. This has interesting policy implications to the extent that trade in intermediates is identified as one of the channels for international transmission of technology with positive effects on the economic growth of integrated economies (Rivera-Batiz and Romer, 1991; Grossman and Helpman, 1991)[5].

The remainder of the paper is organized as follows. Section 1 briefly discusses the standard HOV model and its version amended for failures of factor price equalization and international trade in inputs. Section 2 compares the imputation methodology followed in the construction of the input-output tables used in RTZ studies to that adopted in the compilation of the AIO tables. Section 3 presents the empirical analysis. Section 4 concludes.

1 Theory

1.1 The Standard HOV Model

The original HOV model considers a world with $i, j = 1, \ldots, N$ countries, and at least as many goods $g = 1, \ldots, G$ as factors $f = 1, \ldots, K$. Consumer preferences are identical and homothetic. Each good is produced with constant returns to scale in perfect competition and traded for final as well as intermediate use. Factor markets are competitive. Technologies for all goods and the quality of all factors are identical across countries. There are no trade costs and the distribution of factor endowments is such that factors prices are equalized in equilibrium. Commodity prices are also equalized in equilibrium

[5]Empirical studies investigating the link between imported intermediates and a country's aggregate productivity include Coe and Helpman (1995), Coe et al. (1997), and Xu and Wang (1999). A more recent literature examines the effect of imported inputs on plant-level productivity (Amiti and Konings, 2007; Kasahara and Rodrigue, 2008; Halpern et al., 2006).

as a result of perfectly competitive markets, identical technologies and free trade.

Under these assumptions, the (GxG) input-output matrix, B, is common to all countries. Further, countries use the same average amount of factor f directly to produce one unit of good g, i.e., they are characterized by the same (KxG) technology matrix, D. Hence, techniques of production are identical internationally. Let T_i be country i's vector of net exports and $F_i \equiv D(I-B)^{-1}T_i$ the vector of total factor services embodied in country i's net trade. Noticing that country i's factor content of trade equals the difference between the factor contents of country i's production and consumption, the model delivers the Vanek proposition:

$$F_i = V_i - s_i V^w \iff F_{fi} = V_{fi} - s_i \sum_{j=1}^{N} V_{fj} \qquad \forall i = 1, ..., N \qquad f = 1, ..., K \tag{1}$$

where V_i and V^w are, respectively, country i's and the world vectors of factor endowments, and s_i is country i's share in world spending. Accordingly, a country is a net exporter of the services of those factors with which it is relatively well-endowed (i.e., $F_{fi} > 0 \iff V_{fi}/\sum_{j=1}^{N} V_{fj} > s_i$) and a net importer of the services of those factors with which it is relatively poorly-endowed.

For empirical purposes, knowing the technology matrix and aggregate input-output structure of one economy is sufficient to measure the amount of factor services embodied in any country's net exports. U.S. production techniques are used to estimate this model in section 3.2.

1.2 The HOV Model and Failures of Factor Price Equalization

Studies examining the Vanek prediction when factor prices are not equalized internationally and intermediate goods are traded appeared only recently (Davis and Weinstein, 2001; Hakura, 2001; Trefler and Zhu, 2000, 2005, 2006; Reimer, 2006). Reimer (2006) and Trefler and Zhu (2005, 2006) are the first to provide the definition of factor content of trade that, in these contexts, correctly measures the amount of factors employed worldwide to produce each country's net exports. This definition weights goods flows, whether final or intermediate, according to the technology of the producing country while reconstructing the entire chain of production of traded goods. More formally, the factor content of country i's net exports is given by[6]:

$$F_i = D(I - B)^{-1} T_i \tag{2}$$

where B is an (NGxNG) matrix with its elements $B_{ij}(g, h)$ being the amount of good g used for production of one unit of good h, where g is produced by country i and h by country j, for all i and j; D is the (KxNG) matrix obtained concatenating countries' direct unit input requirement matrices D_i, and T_i is the (NGx1) vector of country i's trade with imports detailed by country of origin.

Trefler and Zhu (2005, 2006) show that under failures of factor price equalization with traded intermediates, the Vanek proposition is valid if each country's consumption of any other country's good is a fixed proportion of the world consumption for that good[7]. This result validates empirical tests of the Vanek proposition when the HOV model accounts for intermediates and failures of factor price equalization.

2 Empirical Challenges

The estimation of the HOV model with international differences in production techniques and global production sharing requires the researcher to obtain data on the source of intermediate inputs. Reimer (2006), and Trefler and Zhu (2005) begin with input-output data taken from the GTAP consortium. In a few cases, national input-output tables distinguish domestic versus foreign sources of inputs, but do not further distinguish the particular source countries from which foreign inputs are purchased. In most cases, however, no sourcing information is provided in the original data. In this case, contributors to the GTAP database separate foreign from domestic sourcing by combining information on countries' aggregate usage of intermediates, trade data and a proportionality assumption

[6]Trefler and Zhu (2005, 2006) develop a thorough proof of this result.

[7]A Vanek consistent model is given by the standard HOV model that allows the distribution of endowments to be such that factor prices are not equalized for any pair of countries. This assumption implies that countries produce in different cones of diversification and complete specialization occurs. Hence, a country's consumption absorption for any good is proportional to the world consumption absorption for that good. Such a model is implicitly considered by Reimer (2006).

in the spirit of Feenstra and Hanson (1996, 1999), and Hummels et al. (2001). Specifically, a country is assumed to import intermediates of a particular good in proportion to the share of imports in the destination country's total use of that good.

More formally, the amount of imported good g used for production of one unit of country i's good h, $B_i^*(g,h)$, is imputed as follows:

$$B_i^*(g,h) = \frac{M_i(g)}{Q_i(g) + M_i(g) - X_i(g)} * \bar{B}_i(g,h) \tag{3}$$

where $\bar{B}_i(g,h)$ is the amount of good g used to produce one unit of country i's good h as obtained from i's national input-output table; $Q_i(g)$, $M_i(g)$ and $X_i(g)$ are country i's total output, total imports and total exports of good g, respectively.

Reimer (2006), and Trefler and Zhu (2005) use GTAP input-output data and extend the proportionality assumption to derive bilateral details[8]. Accordingly, the amount of good g used for production of one unit of good h, where g is produced by country j and h by country i, $B_{ji}(g,h)$, is imputed as follows:

$$B_{ji}^{RTZ}(g,h) = \frac{M_{ij}(g)}{Q_i(g) + M_i(g) - X_i(g)} * \bar{B}_i(g,h) \qquad j \neq i \tag{4}$$

The superscript "RTZ" reflects that the approach outlined in equation (4) underlies the input-output coefficients used by Reimer (2006), and Trefler and Zhu (2005)[9]. The amount of domestic good g used for production of one unit of country i's good h is derived subtracting the per unit usage of imported intermediates from the aggregate per unit usage of inputs and it equals:

$$B_{ii}^{RTZ}(g,h) = \frac{Q_i(g) - X_i(g)}{Q_i(g) + M_i(g) - X_i(g)} * \bar{B}_i(g,h) \tag{5}$$

A problem with the proportionality assumption underlying RTZ data is that it provides accurate input-output coefficients only if international flows of goods respond to

[8]Reimer (2006) actually uses only the U.S. and the rest of the world in his analysis. Thus, he does not need to distinguish U.S. usage of imported intermediates by country of origin.

[9]This imputation methodology provides a more accurate identification of domestic and imported intermediates the more disaggregated are the trade data underlying the calculations. In general, the degree of detail that characterizes the trade data underlying the imputations varies widely across countries and the estimated amount of imports that are imported inputs tend to be downward biased.

trade determinants independent of their end-use. For instance, suppose that U.S. total usage of auto industry products, ai, is \$200 ml (i.e., $Q_{us}(ai)+M_{us}(ai)-X_{us}(ai)$ =\$200 ml) and U.S. imports from the Japanese auto industry amount to \$60 ml (i.e., $M_{us,jap}(ai)$ =\$60 ml). Then, equation (4) predicts that 30% of total auto industry intermediates used into the production of U.S. output, in any industry, are from Japan. This prediction does not change whether the \$60 ml of U.S. imports from Japan are all products for final consumption, all auto parts and components or, half and half. Hence, equation (4) treats Japanese comparative advantage to be independent of the auto industry products end-use. This assumption is problematic if intermediate inputs are produced using different factor intensities than final goods or if input trade is more (or less) sensitive to trade costs than final goods trade.

An alternative way to capture the implications of the proportionality assumption made by RTZ is to look at the per unit usage of imported intermediates relative to the per unit usage of domestic inputs. More formally, country i's relative usage of country j's input g is easily calculated taking the ratio of equations (4) and (5), and it equals the share of good g bilateral imports into domestic production (net of exports) for that good:

$$\frac{B_{ji}^{RTZ}(g,h)}{B_{ii}^{RTZ}(g,h)} = \frac{M_{ij}(g)}{Q_i(g) - X_i(g)} \tag{6}$$

According to RTZ methodology, the relative usage of imported intermediates depends on total bilateral imports of good g and it is independent of the using industry h. The incentive to use imported or domestic products is thus independent of end-use.

This paper exploits the AIO tables as the main data source of detailed input-output structures for nine East Asian countries and the U.S.[10]. The great advantage of the data in the AIO tables is that, domestic and international transactions in intermediates and final goods are distinguished on the basis of country-specific import surveys. More formally, these surveys provide a measure of countries' import matrix, with elements $B_i^*(g,h)$, not based on imputations.

China, Indonesia, Malaysia, Singapore and Thailand conduct complementary surveys

[10]The nine East Asian countries are: China, Indonesia, Japan, Korea, Malaysia, Philippines, Singapore, Taiwan and Thailand.

to obtain the bilateral details of their import input-output structures, $B_{ji}(g,h)$. The remaining countries in the sample identify the origin of imported inputs employing a proportionality assumption according to which a country imports intermediates of a particular good from a given source country in proportion to the share of bilateral imports in the destination country's total imports of that good. Formally, the amount of good g used for production of one unit of good h, where g is produced by country j and h by country i is the imputed as follows:

$$B_{ji}^{AIO}(g,h) = \frac{M_{ij}(g)}{M_i(g)} * B_i^{*,survey}(g,h) \qquad j \neq i \tag{7}$$

where $B_i^{*,survey}(g,h)$ is the amount of imported good g used to produce one unit of country i's good h obtained from that country's survey. Often the imputations in (7) are based on highly disaggregated bilateral trade data. These estimates are then adjusted using additional information on the source country of imported intermediates, and suggestions from local specialists.

The extensive use of country-specific surveys in the AIO tables provides better information on the end-use of traded commodities and can result in pronounced differences relative to the RTZ imputation methodology in two cases. First, many sectors such as the auto or electronics industries are broadly defined so as to encompass both final and intermediate inputs. Suppose the U.S. imports equal values of auto parts from Korea, and final cars from Japan. In any U.S. sector that uses auto industry products as an input, the RTZ method will impute equal input-output coefficients for Japan and Korea. This will overstate Japan's contribution and understate Korea's. Meanwhile, the AIO data will register a large input coefficient for Korea and a zero value for Japan. Applying the same logic to equation (6), the RTZ method will overstate the U.S. relative usage of Japanese inputs and understate that of Korean intermediates.

Second, two different industries may both employ imported electronics inputs but from different sources. Suppose the US auto industry uses Japanese electronics inputs while the U.S. computer industry uses Korean electronics inputs. The RTZ methodology cannot distinguish end-use; rather it notes only the distribution of total electronics imports and so assigns positive coefficients to both Japanese and Korean electronics for both the U.S.

auto and computer industries. Meanwhile, the AIO data will assign zero input coefficients when an industry does not actually employ inputs from that source. The RTZ method will overstate the U.S. relative usage of both Japanese electronics into the computer industry and Korean electronics in the auto industry.

3 Empirics

In this section I compare the AIO data for 2000 to bilateral input-output data for the same set of countries constructed using the RTZ imputation procedure.

In section 3.1 I analyze deviations between input-output coefficients from the two sources to assess the RTZ proportionality assumption. I focus on the prevalence of zero values, the magnitude of differences between AIO and RTZ samples, and the correlation of these differences with known determinants of trade patterns.

Section 3.2 compares the performance of the Vanek proposition both under the standard HOV assumptions and when the model is amended to allow for factor price equalization failure and traded intermediates. To assess how failures of the symmetry assumptions affect the performance of the Vanek proposition, the second model is estimated using the AIO sample and the RTZ sample in turn.

Section 3.3 provides a decomposition of the measured factor content of trade when factor price equalization fails. This decomposition distinguishes the domestic from the foreign nature of factor services embodied in a country's net exports and it is implemented using separately the AIO and the RTZ data. The comparison of the results obtained using the AIO input-output coefficients those based on the RTZ data provides insights into the specific biases introduced by failures of the symmetry hypothesis.

3.1 Empirical Analysis of the Symmetry Hypothesis

3.1.1 Agreement of Distributions

In this section I compare the input use coefficients, $B_{ji}(g,h)$, taken from the AIO and from constructed RTZ data. Recalling equation (4), these coefficients correspond to the use of good g produced by exporting country j, used in industry h within importing country

i, i.e., how much steel produced in Korea is employed in one unit of auto production in Japan. When domestic industry is the source of inputs, the input-output coefficients are denoted $B_{ii}(g,h)$. Each $ji(g,h)$ corresponds to an observation, so for 11 exporters, 11 importers, and 34 industries there are 139876 total values.

Table 1 summarizes the percentage of zero values in both the AIO and the RTZ data. For domestic use of domestic inputs, roughly 20 percent of observations are zeros in both data samples. Considering only the imported input coefficients, there is a dramatically higher incidence of zeros in the AIO data than in the RTZ data. This is precisely what one would expect if the proportionality assumption in the RTZ imputation wrongly conflates flows of final goods and intermediate inputs. Also, this finding is consistent with industrial demand being more specialized than consumer demand. Consumers, indulging their love of variety, demand goods from many sources while firms buy customized inputs from a much narrower set of suppliers.

Next I compare the bilateral distribution of coefficients from the two distributions net of the importing country's sectoral use of given inputs, $\frac{B_{ji}(g,h)}{\bar{B}_i(g,h)}$. Figure 1 scatters these values. If RTZ and AIO data matched, points would line up on the 45 degree line. Clearly, this is not the case. The correlation coefficient is, in both domestic and import samples, about 0.64[11].

For each unit input requirement $B_{ji}(g,h)$, I also look at the ratio $\frac{B_{ji}^{AIO}(g,h)}{B_{ji}^{RTZ}(g,h)}$. Figure 2 shows the histograms for this ratio for domestic and imported inputs[12]. Both distributions show large departures from a value of 1 (AIO and RTZ coefficients equal), but there is a much larger variance for imported inputs. Interestingly, in the import sample, conditional on $\frac{B_{ji}^{AIO}(g,h)}{B_{ji}^{RTZ}(g,h)} < 1$, the mean value equals 0.25; and conditional on $\frac{B_{ji}^{AIO}(g,h)}{B_{ji}^{RTZ}(g,h)} > 1$, the mean value equals 12. These patterns are consistent with countries concentrating purchases of intermediates in a few suppliers[13].

Of course, these ratios could depart wildly from 1 and be of little consequence if an in-

[11]This value is much lower than the one I obtain comparing directly the AIO and RTZ input-output coefficients. An explanation for this finding is that there is a lot of variation across $\bar{B}_i(g,h)$ which inflates the correlation between the distribution of input-output coefficients. In fact, both AIO and RTZ bilateral sourcing details are just shares of the aggregate per unit usage of inputs, $\bar{B}_i(g,h)$, whose value is the same across datasets

[12]Findings are robust to the log transformation of the data.

[13]Hummels and Puzzello (2008) report similar evidence.

put has a minuscule share of gross output. To see whether these differences are significant relative to total production figure 3 plots the difference between the levels of AIO and RTZ unit input requirements, $B^{AIO} - B^{RTZ}$, against the corresponding value in the AIO sample, B^{AIO}. A dashed line identifies the AIO sample mean. Substantial differences between distributions are readily apparent. The percentage of observations for which the difference is zero is about 22% (35%) of the domestic (import) sample, but for only 2% (0.01%) of these observations the ratio between AIO and RTZ values exists and equals 1. Appendix A further analyzes differences in distributions by reporting percentage differences and limits of agreements for the unit input requirements of imported and domestic intermediates. In any case, differences across AIO and RTZ unit input requirements of intermediates are large and relatively more important for imported inputs.

Table 2 provides summary statistics for the AIO and the RTZ relative usages of imported intermediates, $\frac{B_{ji}(g,h)}{B_{ii}(g,h)}$. The number of observations is smaller in the AIO dataset suggesting that RTZ data often overstate countries' use of domestic intermediates. A glance to the means corroborates this finding. The relative demand of imported inputs is, in fact, much smaller in the RTZ dataset than it is in the AIO data. The large discrepancy in the variances of the two distributions is easily explained looking at equation (6). According to RTZ the relative demand of imported inputs is independent of end-use, this naturally reduces the number of values the ratio can take. The existence of large differences between AIO and RTZ relative usages of imported intermediates is also born out by the histogram of their ratio in figure 4.

3.1.2 Formal Test

If differences between AIO and RTZ distributions are not "noise", countries input-output structures based on RTZ method are not an accurate approximation of the true data. In section 2 I argue that differences in the input-output coefficients across datasets would arise because countries' comparative advantage within a product category varies depending on end-use. Now, suppose one was interested in understanding how trade and factor costs (among other sources of comparative advantage) affect countries relative demand of imported inputs. If the AIO and RTZ input-output data were substitutable using one data or the other would not produce significantly different estimated effects. Recalling that RTZ relative usages of imported intermediates capture patterns of total trade while AIO data better identify international flows of inputs, this result would suggest that each explanatory variable has the same effect on intermediate and final goods trade. Suppose, instead, that trade costs have a greater negative effect on AIO rather than RTZ relative demand of imported inputs, this would only be possible if trade in intermediates were more sensitive to trade costs than final goods trade. The proportionality assumption made by RTZ would then be problematic. I follow exactly this identification strategy to test the symmetry hypothesis.

Further, testing for differences between AIO and RTZ relative usages of imported intermediates makes it possible to clearly link empirical failures of the symmetry hypothesis to the biases these failures generate on measured factor contents of trade. When factor price equalization fails the higher is a country's use of imported inputs the more similar are its techniques of production to the foreign ones. Detecting systematic differences in the relative usage of imported intermediates between AIO and RTZ data is then informative of systematic differences in measured factor contents of trade depending on the data used.

Formally, the symmetry hypothesis is tested estimating the following model, for AIO and RTZ countries' relative usage of imported intermediates, separately and taking their ratio, i.e., for $Z \in \left(\frac{B_{ji}^{AIO}(g,h)}{B_{ii}^{AIO}(g,h)}; \frac{B_{ji}^{RTZ}(g,h)}{B_{ii}^{RTZ}(g,h)}; \frac{B_{ji}^{AIO}(g,h)/B_{ii}^{AIO}(g,h)}{B_{ji}^{RTZ}(g,h)/B_{ii}^{RTZ}(g,h)} \right)$:

$$\begin{aligned} logZ = \delta_0 + \delta_1 log(GDP_i) + \delta_2 log\left(\frac{K}{L}\right)_i + \delta_3 log\left(\frac{a_k}{a_l}\right)_{jg} + \delta_4 log\left(\frac{K}{L}\right)_i * log\left(\frac{a_k}{a_l}\right)_{jg} + \\ + \delta_5 log(distance_{ij}) + \delta_6 Contig_{ij} + \delta_7 Lang_{ij} + \delta_8 log(rem_i) + \delta_9 ROW^{imp} + \\ + \phi_j + \phi_{g,h} + \upsilon_{ij}^{g,h} \end{aligned} \tag{8}$$

where GDP_i and $\left(\frac{K}{L}\right)_i$ are the gdp and relative factor endowment of the importing country; $\left(\frac{a_k}{a_l}\right)_{jg}$ is the direct capital-labor intensity of the intermediate g imported from country j; $Contig$ and $Lang$ are indicator variables that take 1 if the trading partners share a border or a primary language, respectively; rem_i is the remoteness index for the importing country. ROW^{imp} is a dummy that takes the value of one if the rest of the world is the importer in the trade partners pair. This indicator variable is introduced to account for measurement error in the data for the ROW. α_j and $\alpha_{g,h}$ are source country, and source-destination sectors fixed effects.

Equation (8) exploits the intuition that the demand of imported intermediates relative to domestic inputs depends upon scale effects as captured by the size of the trade partners, geographical factors and differences in factor costs. Larger countries are the places where agglomeration of production is more likely to take place (Krugman and Venables, 1995, 1996) and should rely relatively more on domestic intermediates. I thus expect a negative estimate for the coefficient of the GDP variable. More remote countries are likely to face high communication, transportation or time barriers which hinder their ability to participate in the international fragmentation of production (Jones and Kierzkowski, 2001, 2003; Harris, 2001) and reduce their use of imported intermediates. Finally, in accordance with the factor proportions theory the global production process should be split to take advantage of international factor cost differentials. Accordingly, a capital abundant country would use relatively more imported intermediates the higher is the labor intensity of required inputs. These products are most likely produced at lowest cost by labor abundant countries[14].

Estimation results for specification (8) are shown in the first three columns of table 4[15]. The symmetry assumption is not supported by the data. In fact, the estimates in the third column of the table show that differences in the relative demand of imported inputs are not just "noise", they are significantly explained by trade determinants[16].

As expected, agglomeration forces matter for the relative demand of imported intermediates: larger and more remote countries rely less on imported intermediates for

[14] Consistently with the literature, I find that countries' techniques of production are biased toward the use of own abundant factors. Estimating the following model for the factor intensity of country j's good g: $log\left(\frac{a_k}{a_l}\right)_{gj} = \gamma_0 + \alpha_g + \gamma_1 \left(\frac{K}{L}\right)_j + \varepsilon_{gj}$, where α_g is a sector fixed-effect, the estimate for γ_1 is positive and strongly significant. This is so whether the direct or total factor intensities of produced goods are considered as the dependent variable. Estimation results are available upon request.

[15] Table 3 reports the sample summary statics.

[16] A joint comparison analysis for aggregate unit input requirement of intermediates, $B_i^*(g,h)$, and for the bilateral import share of intermediates, $\frac{B_{ji}(g,h)}{B_i^*(g,h)}$, shows that the AIO country-specific surveys induce adjustments in both countries' aggregate import matrices and their bilateral details, with the former effect being larger. The empirical failures of the symmetry hypothesis shown in table 4 are consistent with these unreported results.

production; a country's relative demand of foreign inputs is larger with partners with which it shares a border or a primary language. Both effects are stronger when the AIO data are used for the estimation, equivalently, RTZ data understate agglomeration effects. This is exactly what one would expect if trade in intermediates is more sensitive than trade in final goods to trade costs. The only evidence against this conclusion is brought by the distance coefficient estimates.

International factor costs differentials determine the geography of production fragmentation and the AIO data find a stronger countries' tendency to rely relatively more on imported intermediates for inputs intensive in their scarce factor. Differences in factor costs are relatively more important for trade in inputs. As shown in the last column of table 4, this result is robust when the total rather than direct factor intensity is used as an explanatory variable in specification (8)[17].

Additional evidence on the different sensitivity of trade flows to trade determinants depending on end-use is provided by the discrepancy in the R^2 of the models reported in the first two columns of table 4. When RTZ data are used in the estimation of the relative demand of foreign inputs, the explanatory power of the model increases. This is at least suggestive of the fact that the gravity-type variables included in specification (8) explain patterns of trade in final goods better than those in intermediates.

Table 5 shows that results hold when the ROW's input-output coefficients, calculated using mostly data from the GTAP input-output tables, are dropped from the sample[18].

Even though the symmetry hypothesis does fail, the net effect of its empirical failures on the measurement of countries' factor content of trade proves difficult to disentangle. On the one hand, RTZ data, by understating agglomeration effects, would make the techniques of productions for large and more remote countries less intensive in their abundant factors. For these countries the gap in the total capital-labor intensities with other coun-

[17]I use the total factor intensity for this robustness check even though the capital-labor intensity of an exchanged intermediate most likely lies somewhere in between its direct and total capital-labor intensity.

[18]Table 3 shows the sample summary statics. I also verified the robustness of estimates in tables 4 and 5 to the exclusion of those observations for which the demand of imported intermediates relative to domestic inputs is zero using the AIO data but positive according to the RTZ data. The symmetry assumption fails also according to the tobit results even though only robust standard errors could be obtained. The tobit effects at the mean values are significantly higher than the corresponding OLS estimates except for the distance variable which switches sign. Results are available upon request.

tries would be understated and international trade in factor services depressed[19]. The opposite would be true for small or less remote countries. On the other hand, by systematically understating the degree in which each country exchanges intermediates intensive in its scarce factor, RTZ data might inflate the gap between partners' total capital-labor intensities and tend to overstate international trade in factor services.

Nevertheless, the symmetry assumption bias can be quantified comparing trade in factor services as measured estimating the HOV model without factor price equalization using the AIO and the RTZ series in turn.

3.2 Testing the Vanek Proposition

In this section I compare the performance of the Vanek proposition both under the standard HOV assumptions and when the model is amended to allow for factor price equalization failure and traded intermediates. To assess how failures of the symmetry assumptions affect the performance of the Vanek proposition, the second model is estimated using the AIO sample and the RTZ sample in turn.

In order to assess the empirical performance of the Vanek proposition in each model I use empirical tests that are well-established in the literature. First, I perform sign and rank tests in the spirit of Bowen et al. (1987). The sign test asks whether the measured and the predicted factor content match in sign, i.e., $\text{sign}(F_{fi}) = \text{sign}(V_{fi} - s_i \sum_{j=1}^{N} V_{fj})$. The rank test instead checks whether if the amount of services embodied in a country's net export of a factor exceeds that of a second factor the relative abundance of the first factor exceeds that of the second factor, i.e., $F_{fi} > F_{f'i} \Longleftrightarrow V_{fi} - s_i \sum_{j=1}^{N} V_{fj} > V_{f'i} - s_i \sum_{j=1}^{N} V_{f'j}$ with $i = 1, \ldots, N$; $f, f' = K, L$. I gauge the correlation between the measured and the predicted factor content of trade through the Spearman rank correlation coefficient. Further, I implement standard regression tests. Accordingly, I focus on the estimated slope and explanatory power of the linear regression model which explains the measured factor content of trade using the predicted one. The slope estimate and the R^2 from this regression are sufficient to compute the Missing Trade statistic (MT; Trefler, 1995) which is defined as the variance of the measured factor content of trade divided by

[19]This effect is moderated through countries' exports of intermediates, which are likely to be more important for large countries.

the variance of the predicted factor content of trade.

Both the measured and predicted factor content of trade are scaled by a factor-country specific scalar σ_{fi} which allows to express factors in comparable units as well as to account for country size (Trefler, 1995). This strategy ensures that country-factor specific deviations of the measured factor content of trade from the predicted one have unit-variance[20].

Tests results on the performance of the Vanek proposition, for each of the three models, are reported in table 6. Table 7 reports the empirical performance of the Vanek proposition by factor, for each treatment.

The Vanek proposition under the standard HOV model fares poorly, even after having acknowledged the small sample size. With 11 countries and two factors, only twenty-two data points are available to test the models. Measured and predicted factor content of trade match in sign about 60% of the times but the Fisher's exact test suggests that the signs of the factor contents are independent. The rank test is satisfied 45% of the times and the Spearman correlation between measured and predicted factor content of trade is positive but not significantly different from zero. The estimated slope in the regression of the measured on the predicted factor content of trade is close to zero and insignificant. The fact that the variance of the predicted factor content of trade is about 30 times larger than that of the measured just confirms that trade in factor services is largely overstated by the theory[21].

The performance of the Vanek proposition improves in presence of factor price equalization failures and trade in inputs, with the AIO input-output coefficients used for the model estimation. Even though the sign test is satisfied 64% of the times, the hypothesis of independence between the signs of the factor contents of trade cannot be rejected. The rank tests results improve substantially. Factor ranks are correct 91% of the times and the Spearman correlation between the measured and the factor content of trade is high and significant at 0.65. The regression test records a slope significantly different from zero. Trade is still missing, but the variance of the predicted factor content of trade is now "only" ten times that of the measured. These results are consistent with past studies

[20] I follow Trefler (1995) and define $\sigma_{fi} \equiv \sigma_f s_i^{\mu}$, with σ_f being the cross-country standard deviation of $F_{fi} - V_{fi} + s_i \sum_{j=1}^{N} V_{fj}$ and $\mu = 0.9$.

[21] The results in the first row of table 7 reveal a particularly poor performance for labor.

and especially close to Trefler and Zhu (2005) findings[22].

The Vanek proposition performs better when the RTZ rather than the AIO input-output coefficients are used. The sign test is now satisfied 68% of the time at almost 10% significance level[23]. Regression tests improve implying a small decrease in the gap between the variance of the predicted and measured factor contents[24]. The net effect of the symmetry hypothesis failures on the Vanek proposition performance is, thus, positive but small.

This evidence is not surprising in light of the patterns revealed by the last three columns of table 8 which respectively report: the percentage difference in measured factor contents of trade, $\frac{F_{fi}^{AIO}-F_{fi}^{RTZ}}{F_{fi}^{AIO}}$; the signs of countries' measured factor content of trade obtained from the AIO data; and the signs of countries' predicted factor content of trade[25]. The signs of the measured factor content of trade are the same irrespective of the input-output data used in the estimation with the exception of Taiwan for labor. Consistently with the sign tests reported in table 6, in 14 cases out of 22 both AIO and RTZ measured factor contents of trade match in sign the Vanek prediction. Further, the AIO measured factor content of trade tends to be, in absolute value, smaller than the one calculated with RTZ data[26]. According to table 8, Singapore is correctly measured to be a net exporter of the services of capital and the RTZ data overestimate Singapore net trade in capital services by 4.4 percentage points. By inflating trade in factor services most of the times measured and predicted factor content of trade match, RTZ input-output coefficients favor the Vanek proposition[27]. Finally, the symmetry hypothesis bias on the Vanek proposition is small because differences between measured factor contents of trade are small and range only between 0.3 and 8 percentage points.

[22]The Vanek proposition does significantly better under the relaxed assumptions than under the standard HOV framework even when one looks at factor level (see table 7).

[23]The better performance of the sign test is driven by improvements on the sign tests for labor. The sign between measured and predicted labor contents of trade matches 91% of the times. When the AIO data are used this percentage decreases to 82%.

[24]The better performance of the Vanek proposition holds also at factor level (see table 7).

[25]The normalization of the factor contents of trade by σ_{fi} is dropped. The Vanek prediction is independent of the data used in the estimation.

[26]For these calculations the normalization of the factor content of trade by σ_{fi} is dropped. That implies that the Vanek prediction is the same irrespective of the data used in the estimation.

[27]From table 8 Taiwan emerges as an important observation. Dropping the observations for Taiwan does not change the fact that the Vanek proposition performs better when RTZ input-output data are used in the estimation of the amended HOV model.

3.3 Symmetry Hypothesis Failures and Measurement Biases in the Factor Content of Trade

The results in the last two sections are puzzling. On the one hand, the symmetry hypothesis is not supported in the data and AIO input-output coefficients systematically deviate from RTZ ones. On the other hand, adopting the proportionality assumption underlying RTZ data does not produce large differences in countries' measured factor content of trade or in the performance of the Vanek proposition with failures of factor price equalization. In order to clarify this inconsistency I analyze the biases failures of the symmetry hypothesis impose on domestic and foreign components of countries' factor content of trade.

The measured factor content of a country's trade obtained from equation (2) can be easily decomposed into its domestic and foreign components. To see this more clearly consider, without loss of generality, the version of equation (2) with all foreign sources pooled in one aggregate, j:

$$F_i = \begin{bmatrix} D_i & D_j \end{bmatrix} \begin{bmatrix} B^{ii} & B^{ij} \\ B^{ji} & B^{jj} \end{bmatrix} \begin{bmatrix} X_i - M_{ij} \end{bmatrix}$$

where B^{ij} corresponds to the ij-th sub-matrix of $(I-B)^{-1}$ in equation (2). This implies:

$$F_i = \left[D_i(B^{ii}X_i - B^{ij}M_{ij}) + D_j(B^{ji}X_i - B^{jj}M_{ij}) \right] \tag{9}$$

The first term on the right hand side of equation (9) measures the total amount of domestic factors' services embodied into country i's net export (*Domestic Component*). In particular, $D_iB^{ii}X_i$ measures the total amount of domestic factors' services country i contributes to its factor content of trade through the 1st, 2nd, 3rd,... stages of domestic production before exporting; $D_iB^{ij}M_{ij}$ is the total amount of domestic factors' services embodied into the imports from country j. That is, inputs are produced in i using (directly and indirectly) domestic factor services and exported for use in country j's production. When they return as imports from country j they enter with a negative sign in country i's factor content of trade. The second term on the right hand side of equation (9) is the *foreign component* of country i's factor content of trade and it measures the total amount of foreign factors' services embodied into country i's net trade[28]. For convenience, let $(F_i^i + F_i^j)_i$ and $(F_j^i + F_j^j)_i$ represent the domestic and foreign component of country i's measured factor content of trade, respectively.

The (percentage) difference in country i's measured factor content of trade when RTZ input-output coefficients are used in place of the AIO ones can be then decomposed in the (percentage) differences of its components as follows:

$$\frac{F_{fi}^{AIO} - F_{fi}^{RTZ}}{F_{fi}^{AIO}} \equiv \frac{(\sum_i \sum_j (F_{fi}^{j,AIO} - F_{fi}^{j,RTZ}))_i}{F_{fi}^{AIO}} \tag{10}$$

In the data the total amount of domestic factors' services embodied in a country's exports is always greater than the total amount of domestic factors' services embodied in the country's imports (i.e., the domestic component is always positive). Also, the total amount of foreign factors' services embodied in a country's export is always smaller than the total amount of foreign factors' services embodied in the country's imports (i.e., the foreign component is always negative).

Table 8 summarizes the percentage differences between AIO and RTZ data in each component of countries' measured factor content of trade[29]. To interpret the numbers in table 8 it is important to keep in mind the signs of each country's measured factor content of trade and its components; the export part of both domestic and foreign components is positive while their import part is negative.

Interesting patterns emerge from a careful examination of table 8. First, the total amount of domestic factor services embodied in a country's exports is consistently overstated when RTZ data are used. Second, the total amount of both domestic factors' services embodied in a country's imports and foreign factors' services embodied in a country's exports are consistently understated when RTZ data are used[30]. Further, the

[28] The intuition behind the decomposition is easily obtained summing up the factor requirements needed to produce T_i, BT_i, B^2T_i and so forth.

[29] Pooling all foreign sources in a unique aggregate is a relatively innocuous simplification. The empirical tests of the symmetry hypothesis are, in fact, informative of the biases the proportionality assumption underlying RTZ data generate on bilateral trade in factor services.

[30] In the former case, the only exception is Singapore for labor. In the latter case the exceptions are

proportionality assumption underlying RTZ data tend to overstate the total amount of foreign factor services embodied in a country's imports. Finally, the gross quantity of both the domestic and the foreign factor services embodied in a country's trade tend to be overstated by RTZ data. These biases are opposite in sign and tend to cancel each other out in systematic ways: the bias in the domestic component tends to prevail when a country is measured net exporter of a factor's services; the bias in the domestic component tends to prevail when a country is measured net importer of a factor's services.

These patterns are consistent with the proportionality assumption made in RTZ data inflating the gap in total factor intensities across countries[31]. The symmetry hypothesis tend to overstate the use of domestic relative to imported inputs for domestic production, i.e., for given exports and imports, RTZ unit input requirements of intermediates tend to understate the exchange in factor services across countries through trade in intermediates[32]. This inflates both the domestic and foreign components of countries' measured factor content of net exports generating a small positive bias on net flows of factors and in the performance of the Vanek proposition.

4 Conclusions

Davis and Weinstein (2003) in their review of the factor content of trade literature state: "Future work needs to gather better and more extensive data sets, to consider more carefully the role of traded intermediates, cross-countries differences in demand, the role of trade costs, and so on."

This paper succeeds in fulfilling two of the needs of the literature. It exploits the AIO tables whose bilateral input-output structures are derived from extensive use of country-specific surveys. The greater accuracy of input-output coefficients allows a better assessment of the Vanek proposition in the presence of unequal factor prices and global production sharing.

In particular, I show that past studies examining the performance of the Vanek propo-

Japan, U.S. and Malaysia for labor.

[31]Looking at the data, the average gap between partners' total capital-labor intensities calculated on RTZ data is, on average, 0.5% larger than the one based on AIO data.

[32]Notice that this finding is also consistent with the sample statistics discussed in section 3.1.1.

sition under relaxed assumptions on international production techniques might be overoptimistic. This bias arises due to failures of the symmetry hypothesis underlying the data exploited to test the model.

The symmetry assumption is implied by theoretical models widely used in the literature as standard models of monopolistic competition with intermediates (Krugman and Venables, 1996; Hillberry and Hummels, 2002; Redding and Venables, 2004)[33] and the general equilibrium model proposed by Eaton and Kortum (2002).

For instance, the structure of the Eaton and Kortum (2002) model implies that the patterns of trade are the same in intermediates and in final goods. Intuitively, if a country finds it cheaper to buy a manufacturing good from a particular source, it will do so independent of the "end-use" of the good. The implicit assumption is that the cross-country differences in technology and factor costs have the same effect on trade in intermediate and final goods.

The evidence I bring in this paper questions this assumption and points out further research is needed to deepen our understanding of the role comparative advantage forces play in determining specialization patterns.

[33] Hummels and Puzzello (2008) show standard models of monopolistic competition with intermediates imply that the input share of bilateral trade in an industry is not explained by factor and trade costs but only by its industrial absorption share.

References

Altman, D., Bland, J., 1983. Measurement in medicine: the analysis of method comparison studies. Statistician 32(3), 307-317.

Amiti, M., Konings, J., 2007. Trade liberalization, intermediate inputs and productivity: evidence from Indonesia. American Economic Review 97(5), 1611-1638.

Bergstrand, J.H., Egger, P., 2008. The growth of FDI relative to trade. Mimeo.

Bland, J., Altman, D., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet i, 307-310.

Bland, J., Altman, D., 1995. Comparing methods of measurement: why plotting difference against standard methods is misleading. Lancet 346, 1085-1087.

Bowen, H.P., Leamer, E.E., Sveikauskas, L., 1987. Multicountry, multifactor tests of the factor abundance theory. American Economic Review 77(5), 791-809.

Coe, D.T., Helpman, E., 1995. International R&D spillovers. European Economic Review 39, 859-887.

Coe, D.T., Helpman, E., Hoffmaister, A.W., 1997. North-South R&D spillovers. Economic Journal 107, 134-149.

Davis, D.R., Weinstein, D.E., 2001. An account of global factor trade. American Economic Review 91(5), 1423-1453.

Davis, D.R., Weinstein, D.E., 2003. The factor content of trade, in: Choi, E.K., Harrigan, J. (Eds.), Handbook of International Trade, Basil Blackwell, pp. 119-146.

Eaton, J., Kortum, S.S., 2002. Technology, geography and trade. Econometrica 70(5), 1741-1779.

Feenstra, R.C., Hanson, G.H., 1996. Globalization, outsourcing and wage inequality. American Economic Review 86(2), 240-245.

Feenstra, R.C., Hanson, G.H., 1999. The impact of outsourcing and high-technology capital on wages: estimates from the United States, 1979-1990. Quarterly Journal of Economics 114(3), 907-940.

Grossman, G.M., Helpman, E., 1991. Innovation and growth in the global economy. MIT Press.

Hakura, D.S., 2001. Why does HOV fails? The role of technological differences within the EC. Journal of International Economics 54(2), 361-382.

Halpern, L., Koren, M., Szeidl, A., 2006. Imports and productivity. Mimeo, Federal Reserve Bank of New York.

Harris, R.G., 2001. A communication-based model of global production fragmentation, in: Arndt, S.W., Kierzkowski H. (Eds.), Fragmentation: new production patterns in the world economy.

Hillberry, R.H., Hummels, D.L., 2002. Explaining home bias in consumption: the role of intermediate input trade. NBER Working Paper No.9020.

Hummels, D.L., Puzzello, L., 2008. Some evidence on the nature and growth of input trade. Mimeo, Purdue University.

Hummels, D.L., Ishii, J., Yi, K.-M., 2001. The nature and growth of vertical specialization in world trade. Journal of International Economics 54(1), 75-96.

Jones, R.W., Kierzkowski, H., 2001. A framework for fragmentation, in: Cheng, L.K., Kierzkowski, H. (Eds.), Fragmentation and international trade.

Jones, R.W., Kierzkowski, H., 2003. International fragmentation and the new economic geography.

Krugman, P.R., Venables, A.J., 1995. Integration, specialization and adjustment. NBER Working Paper 4559.

Krugman, P.R., Venables, A.J., 1996. Integration, specialization and adjustment. European Economic Review 40(3-5), 959-967.

Kasahara, H., Rodrigue, J.., 2008. Does the use of imported intermediates increase productivity? Plant-level evidence. Journal of Development Economics 87, 106-118.

Leamer, E.E., 1984. Sources of international comparative advantage: theory and evidence. MIT Press.

Leontief, W.W., 1953. Domestic production and foreign trade: the American capital position re-examined. Proceedings of the American Philosophical Society 97(4), 332-349.

Maskus, K.E., 1985. A test of the Heckscher-Ohlin-Vanek theorem: the Leontief commonplace. Journal of International Economics 19(3-4), 201-212.

Redding, S., Venables, A.J., 2004. Economic geography and international inequality. Journal of International Economics 62(1), 53-82.

Reimer, J.J., 2006. Global production sharing and trade in the services of factors. Journal of International Economics 68(2), 384-408.

Repetto, A., Ventura, J., 1998. The Leontief-Trefler hypothesis and factor price insensitivity. Mimeo, Massachusetts Institute of Technology.

Rivera-Batiz, L.A., Romer, P.M., 1991. Economic integration and endogenous growth. Quarterly Journal of Economics 106(2), 531-555.

Romalis, J., 2004. Factor proportions and the structure of commodity trade. American Economic Review 94(1), 67-97.

Schott, P.K., 2003. One size fits all? Heckscher-Ohlin specialization in global production. American Economic Review 93(3), 687-708.

Sitchinava, N., Blonigen, B.A., Davies, R.B., 2009. Contracts, market thickness and outsourcing. Mimeo, University of Oregon.

Trefler, D., 1993. International factor price differences: Leontief was right!. Journal of Political Economy 101(6), 961-987.

Trefler, D., 1995. The case of missing trade and other mysteries. American Economic Review 86(5), 1029-1046.

Trefler, D., Zhu, S.C., 2000. Beyond the algebra of explanation: HOV for the technology age. American Economic Review Papers and Proceedings 90(2), 145-149.

Trefler, D., Zhu, S.C., 2005. The structure of factor content predictions. NBER Working Paper 11221.

Trefler, D., Zhu, S.C., 2006. The structure of factor content predictions. Mimeo, Michigan State University.

Xiang, C., 2007. Diversification cones, trade costs and factor market linkages. Journal of International Economics 71(2), 448-466.

Xu, B., Wang, J., 1999. Capital goods and R&D spillovers in the OECD. Canadian Journal of Economics 32(5), 1258-1274.

Table 1: Percentage of Zeros in RTZ and AIO Samples of Unit Input Requirements

	Domestic (B_{ii},%)	Imported (B_{ji},%)	All
$B^{AIO}(g,h)$	20.51	50.60	47.86
$B^{RTZ}(g,h)$	20.00	34.62	33.29

Table 2: Relative Usage of Imported Intermediates in RTZ and AIO Samples

	% Zeros	Mean (st. dev.)	N
$RUI^{AIO}(g,h)$	38.30	0.1340	101110
		(7.889)	
$RUI^{RTZ}(g,h)$	18.32	0.0704	101750
		(1.0771)	

Note. RUI stands for relative usage of imported intermediates, $\frac{B_{ji}(g,h)}{B_{ii}(g,h)}$.

Table 3: Sample Statistics

Variables	Units	Full sample		w/o ROW	
		Mean	Std. Dev.	Mean	Std. Dev.
RUI^{AIO}		0.218	10.052	0.203	11.521
RUI^{RTZ}		0.084	1.189	0.060	0.271
$\frac{RUI^{AIO}}{RUI^{RTZ}}$		5.820	316.794	6.625	372.196
$log\left(RUI^{AIO}\right)$		-5.652	2.973	-5.618	2.817
$log\left(RUI^{RTZ}\right)$		-5.389	2.840	-5.126	2.417
$log\left(\frac{RUI^{AIO}}{RUI^{RTZ}}\right)$		-0.263	1.667	-0.493	1.671
GDP_i	\$1,000	4.32e+09	5.67e+09	2.07e+09	3.23e+09
$(\frac{K}{L})_i$		79.518	86.989	90.761	91.764
$(\frac{a_k}{a_l})_{jg}^{Direct}$		197.232	832.899	209.362	873.542
$(\frac{a_k}{a_l})_{jg}^{Total}$		75.899	101.783	78.22715	101.469
$distance_{ij}$	kms	6513.90	4424.69	5277.49	4578.90
$Contig$		0.047	0.212	0.065	0.247
$Lang$		0.142	0.350	0.197	0.398
rem_i	kms	8086.54	2305.57	8896.02	1652.73

Note. RUI stands for relative usage of imported intermediates, $\frac{B_{ji}(g,h)}{B_{ii}(g,h)}$.

Table 4: Empirical Results for the Relative Demand of Imported Intermediates

Dependent Variables	$log\Big(RUI^{AIO}\Big)$	$log\Big(RUI^{RTZ}\Big)$	$log\Big(\frac{RUI^{AIO}}{RUI^{RTZ}}\Big)$	$log\Big(\frac{RUI^{AIO}}{RUI^{RTZ}}\Big)$
$log(GDP_i)$	-0.8576***	-0.6259***	-0.2317***	-0.2331***
	(0.0760)	(0.0622)	(0.0293)	(0.0298)
$log((\frac{K}{L})_i)$	0.4450***	0.4309***	0.0141	0.0285
	(0.0860)	(0.0792)	(0.0372)	(0.0445)
$log((\frac{a_k}{a_l})_{jg}^{Direct})$	0.3737***	0.3415***	0.0322	
	(0.0845)	(0.0837)	(0.0347)	
$log((\frac{K}{L})_i) * log((\frac{a_k}{a_l})_{jg}^{Direct})$	-0.0561***	-0.0406**	-0.0154**	
	(0.0212)	(0.0195)	(0.0077)	
$log((\frac{a_k}{a_l})_{jg}^{Total})$				0.0343
				(0.0574)
$log((\frac{K}{L})_i) * log((\frac{a_k}{a_l})_{jg}^{Total})$				-0.0200*
				(0.0103)
$log(distance_{ij})$	-0.2162*	-0.3643***	0.1481***	0.1484***
	(0.1295)	(0.0968)	(0.0468)	(0.0469)
$Contig$	1.0057***	0.6198***	0.3859***	0.3871***
	(0.2691)	(0.2194)	(0.0996)	(0.0994)
$Lang$	0.3873**	0.2309	0.1565**	0.1548**
	(0.1827)	(0.1534)	(0.0625)	(0.0632)
$log(rem_i)$	-1.9799***	-0.5534	-1.4264***	-1.4339***
	(0.7510)	(0.5550)	(0.2911)	(0.2942)
ROW^{imp}	0.9511**	-0.0947	1.0458***	1.0478***
	(0.3683)	(0.2723)	(0.1457)	(0.1466)
$(g,h)fixed-effects$	yes	yes	yes	yes
$jfixed-effects$	yes	yes	yes	yes
R^2	0.4185	0.5949	0.1399	0.1400
N	62271	62271	62271	62271

Note. RUI stands for relative usage of imported intermediates, $\frac{B_{ji}(g,h)}{B_{ii}(g,h)}$. Standard errors clustered at country-pair level are in parenthesis. *:P<0.1; **:P<0.05; ***:P<0.01.

Table 5: Results for the Relative Demand of Imported Intermediates, no ROW

Dependent Variables	$log\left(RUI^{AIO}\right)$	$log\left(RUI^{RTZ}\right)$	$log\left(\frac{RUI^{AIO}}{RUI^{RTZ}}\right)$	$log\left(\frac{RUI^{AIO}}{RUI^{RTZ}}\right)$
$log(GDP_i)$	-0.8576*** (0.0836)	-0.6071*** (0.0706)	-0.2505*** (0.0306)	-0.2506*** (0.0310)
$log((\frac{K}{L})_i)$	0.4791*** (0.0841)	0.4719*** (0.0813)	0.0073 (0.0368)	0.0047 (0.0441)
$log((\frac{a_k}{a_l})_{jg}^{Direct})$	0.3030*** (0.0883)	0.3026*** (0.0840)	0.0004 (0.0343)	
$log((\frac{K}{L})_i) * log((\frac{a_k}{a_l})_{jg}^{Direct})$	-0.0649*** (0.0202)	-0.0501** (0.0192)	-0.0148* (0.0077)	
$log((\frac{a_k}{a_l})_{jg}^{Total})$				-0.0352 (0.0523)
$log((\frac{K}{L})_i) * log((\frac{a_k}{a_l})_{jg}^{Total})$				-0.0148 (0.0103)
$log(distance_{ij})$	0.0025 (0.1023)	-0.2250*** (0.0833)	0.2275*** (0.0425)	0.2266*** (0.0426)
$Contig$	1.3024*** (0.2993)	0.8430*** (0.2321)	0.4594*** (0.1240)	0.4575*** (0.1238)
$Lang$	0.4718** (0.1869)	0.2905* (0.1545)	0.1814*** (0.0642)	0.1813*** (0.0651)
$log(rem_i)$	-1.3150* (0.6784)	-0.0327 (0.5561)	-1.2823*** (0.2512)	-1.2844*** (0.2536)
$(g,h) fixed-effects$	yes	yes	yes	yes
$j fixed-effects$	yes	yes	yes	yes
R^2	0.3940	0.4874	0.0412	0.0410
N	45025	45024	45024	45024

Note. RUI stands for relative usage of imported intermediates, $\frac{B_{ji}(g,h)}{B_{ii}(g,h)}$. Standard errors clustered at country-pair level are in parenthesis. *:P<0.1; **:P<0.05; ***:P<0.01.

Table 6: Empirical Performance of the Vanek Proposition

	Sign Test	Rank Test	Spearman Corr.	Slope	R^2	MT	Obs.
HOV standard	0.5909	0.4545	0.1790	0.0153	0.0070	0.0335	22
	(0.429)		(0.425)	(0.711)			
HOV, No FPE & Traded Inputs	0.6363	0.9091	0.6533	0.1905	0.3839	0.0945	22
(AIO Data)	(0.229)		(0.001)	(0.002)			
HOV, No FPE & Traded Inputs	0.6818	0.9091	0.6544	0.1954	0.3840	0.0994	22
(RTZ Data)	(0.113)		(0.001)	(0.002)			

Note. P-values in parenthesis;

Table 7: Empirical Performance of the Vanek Proposition by Factor

	Capital				Labor		
	Slope	R^2	MT	N	Slope	R^2	MT
HOV standard	0.1948	0.2556	0.1485	11	0.0005	0.0095	0.00003
	(0.113)				(0.776)		
HOV, No FPE & Traded Inputs	0.4123	0.3816	0.4456	11	0.1918	0.8431	0.0436
(AIO Data)	(0.043)				(0.000)		
HOV, No FPE & Traded Inputs	0.4273	0.3838	0.4757	11	0.1962	0.8431	0.0456
(RTZ Data)	(0.042)				(0.000)		

Note. P-values in parenthesis;

Table 8: Decomposition of Percentage Differences in Measured Factor Content of Trade

Panel A. Decomposition of Measured Capital Content of Trade, (%)							
Country	Δ% Domestic Services embodied in		Δ% Foreign Services embodied in		Total % diff.	Sign	Sign
	Exports (+)	Imports (-)	Exports (+)	Imports (-)	MFCT(,%)	MFCT	PFCT
China	-3.65	-0.08	2.43	-0.38	-1.67	+	-
Indonesia	-2.68	-0.04	2.72	-0.70	-0.70	+	-
Japan	-1.56	-0.43	0.53	-0.05	-1.51	+	+
Korea	-16.35	-0.22	10.80	-0.54	-6.31	+	-
Malaysia	-2.59	-0.24	0.93	0.37	-1.53	+	+
Philippines	-7.33	-0.14	7.09	-0.36	-0.73	+	+
Singapore	-8.30	-0.06	4.02	-0.09	-4.43	+	+
Taiwan	-18.08	-0.64	27.73	1.00	10.02	+	-
Thailand	-4.51	-0.05	3.97	-0.18	-0.76	+	+
USA	-4.21	-4.23	2.88	-1.53	-7.09	+	-
ROW	0.35	0.47	-2.34	-1.26	-2.77	-	+
Panel B. Decomposition of Measured Labor Content of Trade, (%)							
China	-2.00	-0.05	0.17	0.01	-1.86	+	+
Indonesia	-3.04	-0.03	0.58	0.03	-2.46	+	+
Japan	0.13	0.06	0.25	-1.02	-0.57	-	-
Korea	9.83	0.10	-16.29	-1.66	-8.02	-	-
Malaysia	-1.67	-0.41	-3.11	2.46	-2.72	+	+
Philippines	-4.77	-0.07	1.52	0.11	-3.20	+	+
Singapore	10.51	-0.04	-2.33	-7.85	0.29	-	-
Taiwan	-103.49	-3.39	213.06	42.15	148.32	**+***	-
Thailand	-3.20	-0.03	1.57	0.23	-1.43	+	+
USA	0.21	0.29	0.36	-0.43	0.43	-	-
ROW	0.63	0.28	-2.07	-2.33	-3.50	-	+

Note. The signs next to the Exports/Imports headings track the sign of the corresponding component in absolute terms. The domestic component is always positive while the foreign component is always negative. *The sign is in bold to identify that Taiwan is measured net exporter of labor services according to the AIO data but a net importer of labor services using RTZ input-output coefficients.

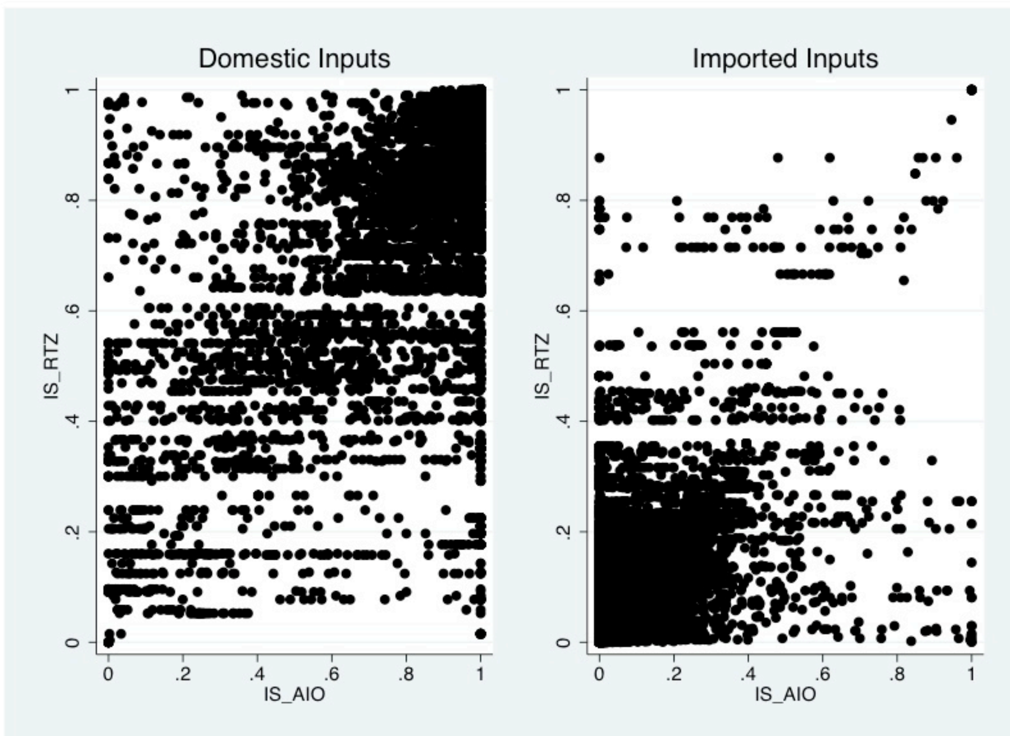

Figure 1: Input Shares: AIO vs. RTZ Data

Note. IS stands for Input Share, $\frac{B_{ji}(g,h)}{B_i(g,h)}$.

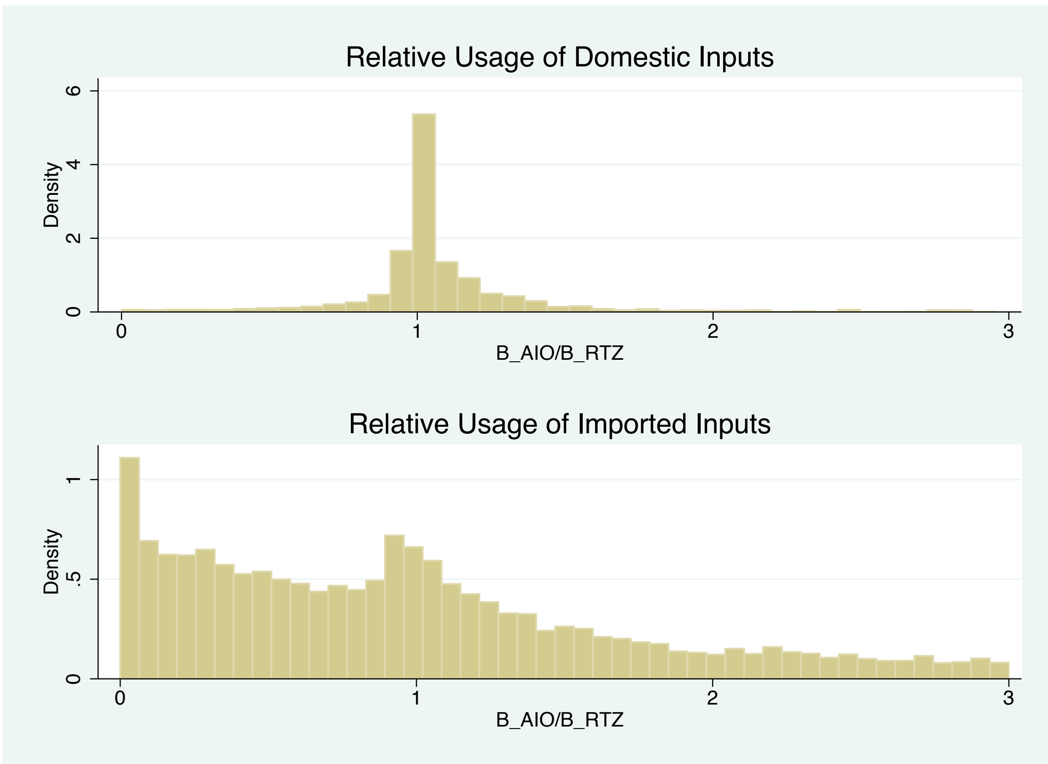

Figure 2: Usage of Inputs: AIO vs. RTZ Data

Note. For legibility of graphs the support is restricted to (0, 3]. This implies dropping about 21% and 34% of the observation in the domestic and import samples, respectively. In both samples, these observations are mostly zeros. In the import sample more than 9% of the distribution extends after 3.

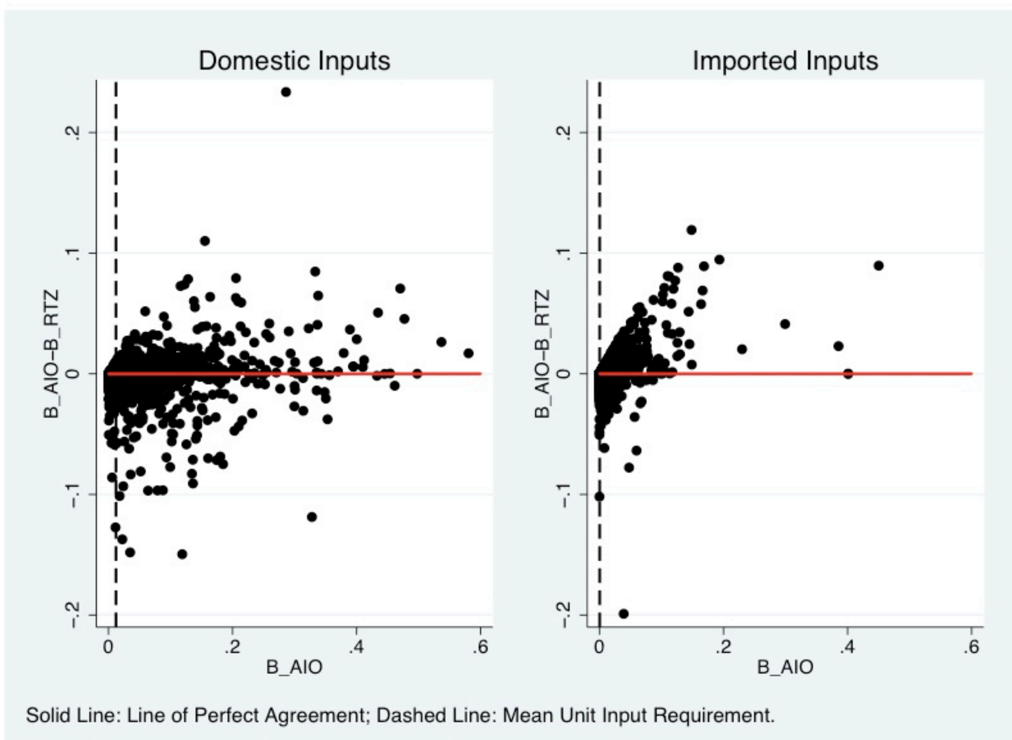

Figure 3: Differences in Usage of Inputs: AIO vs. RTZ Data

Note. The mean unit input requirement of domestic intermediates is 1.2%. The mean unit input requirement of imported intermediates is 0.03%.

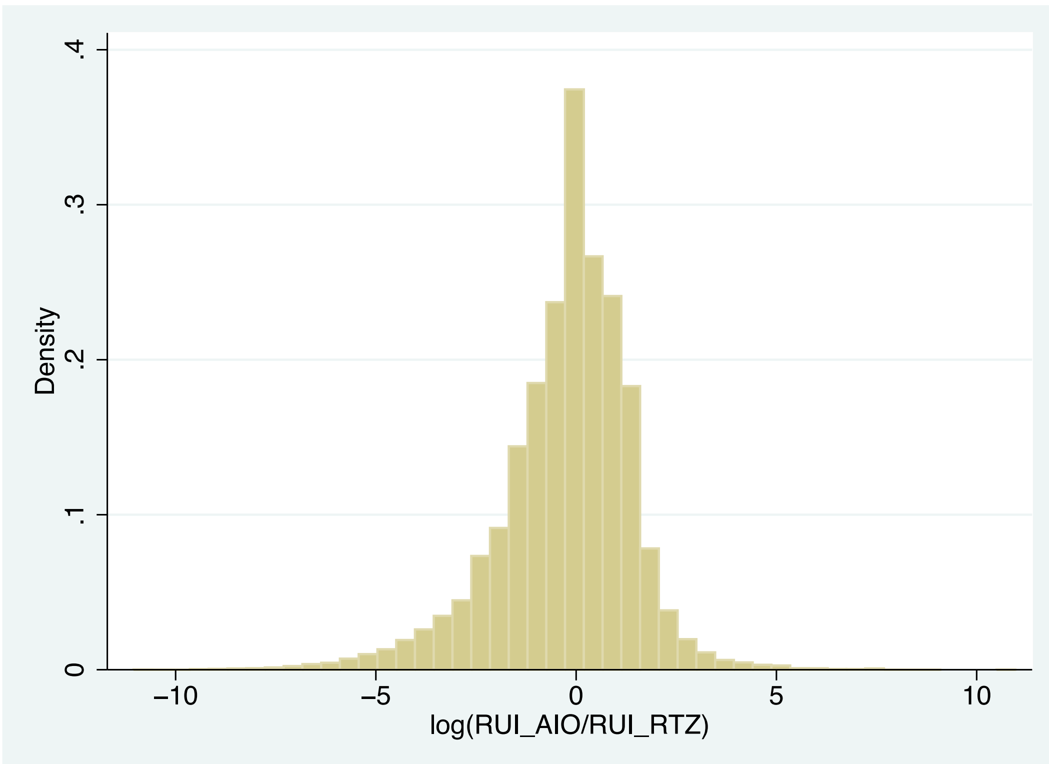

Figure 4: Relative Usage of Imported Inputs: AIO vs. RTZ Data

Note. RUI stands for relative usage of imported intermediates, $\frac{B_{ji}(g,h)}{B_{ii}(g,h)}$.

Appendix A. Differences in Unit Input Requirements

Percentage Differences

Tables A1 and A2 report the percentage differences in unit input requirements of imported and domestic intermediates by destination country and using sector. The numbers in these tables are computed on unit input requirements aggregated across partners and sourcing sectors. Percentage differences are largest among unit input requirements of imported intermediates, and especially large for Thailand.

Limits of Agreement

Here I compare the AIO and RTZ distributions of domestic and imported intermediates using 95% limits of agreement. The concept of limits of agreement has been introduced in the medical literature on comparison methods by Bland and Altman (1983). The idea is to generate a 95% confidence interval for the difference between two measurements. In my particular case, if the RTZ data are within a reasonably narrow neighborhood of the AIO data ninety-five percent of the times, the two measures of unit input requirements are substitutable. The calculation of limits of agreement is straightforward when differences between measures do not significantly correlate with the magnitude of the measurement[34] and follow a normal distribution. To satisfy the first requirement a log transformation is required for the unit input requirements of imported intermediates. Hence, some information is lost. Looking at table A3 it emerges that most of AIO unit input requirements of imported intermediates are between 13.37 times and 0.04 times the corresponding RTZ value. For the domestic case, the limits are relatively narrower with most of the AIO unit input requirements differing from the RTZ by between plus or minus 0.013 units. This analysis is consistent with differences across AIO and RTZ unit input requirements of intermediates being large and relatively more important for imported inputs[35].

[34] I do verify the absence of correlation between differences in AIO and RTZ unit input requirements and their average. Spurious correlation between difference and magnitude could be detected if the AIO value was considered instead of the average (see Bland and Altman, 1995).

[35] The limits of agreement for the relative usages of imported intermediates suggest that 95% of AIO relative usages of imported intermediates are between 20.18 times and 0.03 times the corresponding RTZ values.

Table A1: Percentage Differences in Usages of Imported Inputs by Destination Country and Destination Sector (%)

Sec	Chn	Idn	Jpn	Kor	Mys	Twn	Phl	Sgp	Tha	Usa
1	-139.59	9.26	-28.73	-198.32	-178.9	-691.58	21.01	–	37.46	–
2	-138.82	-89.94	-243.22	-385.16	-88.71	-190.46	17.64	–	-1155.7	-25.27
3	-0.38	-28.46	27.37	12.81	-50.7	18.87	-20	17.27	12.3	-17.24
6	-147.67	-40.62	-53.14	-423.28	-209.9	-998.61	1.63	-47.98	-1717.9	-102.7
7	-28.2	-37.68	-1177.0	-286.17	18.2	-2629.4	1.06	0	-5316.9	4.35
8	-42.72	31.35	-6.82	9.18	-241.34	49.43	-17.08	-14.62	-144.01	-59.17
9	-7.01	-94.49	-114.81	0	31.25	-19.23	4.92	–	-1370.1	-1
10	2.91	-71.33	-245.16	-223.74	-13.82	12.69	4.11	-7.03	-663.51	-8.7
11	-303.39	16.44	32.72	-102.04	-63.83	-185.8	-12.31	-0.35	-22.15	-12.26
13	24.77	26.05	-13.56	-0.26	2.98	-79.91	-50.08	34.82	23.69	14.44
16	-7.7	-1.77	34.77	5.69	28.55	-56.05	28.59	-19.74	11.9	29.82
17	11.82	32.17	-26.39	15.3	5.22	-64.94	52.8	-1.74	-13.56	-34.21
18	-26.42	18.25	-63.54	-32.56	8.05	-16.72	14.55	2.2	-36.11	3.28
19	51.53	-23.53	-0.78	56.15	-45.68	52.32	-41.68	0.66	28.8	-79.21
20	-4.56	5.36	35.65	39.42	-10.27	19.78	29.65	-0.07	-12.97	14.18
21	56.68	36.91	35.03	26.87	20.35	31.5	37.67	13.61	45.54	23.55
22	23.68	52.95	1.69	0.52	8.12	8.07	1.54	14	27.78	6.88
23	12.85	27.08	38.88	33.17	-4.56	19.56	31.67	5.34	12.39	11.74
24	-34.58	-20.6	-119.71	-90.63	-12.16	0.85	9.28	-3.12	5.44	-12.39
25	9.8	2.17	18.51	8.01	18.9	9.29	2.72	12.75	-20.47	1.83
26	14.96	-4.36	53.32	35.4	17.24	10.44	-6.53	-11.88	42.99	0.56
27	-4.1	8.05	-61.85	-30.57	13.61	-57.37	3.73	-32.46	11.69	-8.15
28	-6.62	-20.79	-63.98	-31.72	8.85	-14.92	-4.69	-28.69	16.48	2.7
30	24.35	9.36	8.75	34.41	2.03	16.49	5.76	13.02	9.55	21.92
31	-12.89	24.29	-17.03	-32.3	0.17	-7.3	12.99	8.96	6.71	-2.11
32	14.26	11.7	-8.27	14.08	-8.35	-13.39	-14.2	-13.13	25.96	-0.41
33	-71.22	-133.68	-3.5	4.48	-20.61	-4.56	-1.01	-385.3	-706.83	-7.57
34	-27.93	1.72	-98.44	-194.26	-13.1	-94.86	-3.97	-41.44	-26.75	-59.67
35	-53.47	-36.96	16.11	22.61	39.37	-76.5	-6.81	-91.08	-202.93	-24.52
36	-49.27	0.29	52.04	58.19	6.12	43.63	-6.02	33.79	-504.05	22.67
37	-44.08	-14.93	20.42	-64.95	10.88	31.72	6.65	-30.31	-5354.1	51.08
38	16.33	-11.53	26.32	4.32	37.36	18.29	5.04	45.65	-839.98	39.48
39	-1.72	-132.73	-19.9	-73.53	18.29	-45.22	-5.58	-26.05	-47.25	-30.11
40	-11.91	-22.6	-193.81	1.21	25.36	16.27	-10.27	-11.02	-3106.6	-2.39
μ	-26.48	-13.9	-63.47	-52.58	-18.85	-143.76	2.73	-16.59	-615.97	-7.02
σ	68.67	45.32	211.35	121.67	66.89	486.31	20.33	69.97	1368.8	31.9

Table A2: Percentage Differences in Usages of Domestic Inputs by Destination Country and Destination Sector (%)

Sec	Chn	Idn	Jpn	Kor	Mys	Twn	Phl	Sgp	Tha	Usa
1	4.16	-2.07	0.96	7.43	21.08	11.97	-6.3	0	-19.35	0
2	4.15	11.64	7.1	17.19	61.4	10.7	-4.97	0	21.52	1.49
3	0.02	2.71	-3.23	-2.91	9.73	-6.43	2.72	-21.67	-2.79	1.01
6	1.84	3.15	1.97	8.85	11.19	12.42	-0.32	19.77	11.07	2.85
7	1.97	6.81	9.76	12.62	-13.53	13.11	-0.37	0	33.3	-0.19
8	1.36	-11.27	0.58	-1.8	19.34	-25.21	3.33	8.43	9.64	1.79
9	0.57	10.56	1.62	0	-23.23	11.23	-1.26	0	19.52	0.2
10	-0.24	8.93	3.85	7.23	6.07	-8.7	-1.59	3.1	19.48	0.64
11	1.66	-1.2	-2.45	2.99	15.34	3.62	0.96	0.28	1.57	0.38
13	-0.76	-1.17	1.06	0.01	-0.94	4.31	1.33	-76.55	-2.26	-3.89
16	0.29	0.3	-2.64	-0.72	-15.59	9.03	-5.35	7.94	-3.76	-3.2
17	-1.11	-13.91	2.94	-3.69	-4.57	12.78	-77.44	1.79	3.6	3.94
18	1.92	-4.97	6.35	4.78	-7.14	4.18	-18.51	-2.14	5.89	-0.72
19	-10.47	3.17	0.16	-35.1	20.83	-60.14	18.5	-0.56	-12.8	13.79
20	0.54	-0.72	-6.78	-20.17	1.73	-16.4	-10.2	0.04	6.99	-1.63
21	-13.7	-21.2	-2.1	-5.33	-15.17	-12.19	-32.59	-7.83	-34.57	-2.28
22	-6.41	-37.85	-4.05	-2.47	-3.99	-58.84	-6.38	-140.62	-161.49	-2.7
23	-1.45	-13.22	-4	-9.97	2.63	-10.85	-34.07	-3.98	-5.58	-1.71
24	1.8	3.25	6.04	9.67	4.47	-0.35	-2.96	2.07	-1.47	0.98
25	-0.89	-0.78	-1.06	-1.44	-23.41	-3.45	-2.67	-11.91	12.62	-0.22
26	-1.95	0.66	-33.64	-29.16	-24.01	-5.6	2.82	7.51	-52.01	-0.1
27	0.34	-3.8	2.15	4	-15.83	10.59	-4.26	17.34	-12.7	0.93
28	0.55	4.98	2.05	4.33	-7.58	5.72	3.27	14.57	-13.92	-0.54
30	-6.07	-3.71	-0.96	-23.56	-5.8	-15.66	-27.39	-19.56	-20.31	-6.62
31	1.39	-67.97	1	6.26	-0.25	3.88	-20.82	-10.88	-5.66	0.32
32	-1.69	-2.34	0.68	-3.37	4.72	4.82	6.53	8.31	-18.96	0.06
33	3.64	10.83	0.73	-2.44	5.79	1.23	0.28	49.36	26.48	0.91
34	1.77	-0.49	3.16	9.81	6.25	15.62	1.6	14.86	7.01	3.85
35	2.58	5.12	-0.62	-2.74	-19.83	5.68	1.09	16.04	8.91	0.91
36	1.89	-0.09	-5.91	-38.83	-3.5	-21.87	2.14	-52.16	15.75	-1.84
37	3.62	1.72	-0.61	2.91	-7.22	-6.7	-1.75	7.51	13	-5.61
38	-0.99	1.67	-0.9	-0.28	-9.81	-2.64	-0.9	-22.67	11.77	-1.09
39	0.11	6.92	0.66	4.05	-8.99	4.5	1.24	6.63	6.91	1.18
40	0.57	4.26	2.72	-0.16	-10.17	-5.62	1.95	4.61	9.07	0.16
μ	-0.26	-2.94	-0.39	-2.41	-0.88	-3.39	-6.25	-5.3	-3.63	0.09
σ	3.82	15	6.86	12.91	16.61	17.68	16.64	31.78	32.85	3.33

Table A3: Limits of Agreements by Import Status

	N	Upper Limit	Lower Limit
$B_{ii}(g,h)$	12716	0.013	-0.013
$B_{ji}(g,h)^a$	62726	13.37	0.04

[a]Reported limits of agreement are antilogs of the true intervals.

Appendix B. Data

Brief Data Description

The AIO tables are collected by the Institute of Developing Economies (IDE, henceforth) and are available for the years 1975, 1985, 1990, 1995 and 2000. The paper uses the 2000 AIO tables which report detailed information on intermediate input usage, domestic consumption, gross output, value added and trade data for nine Asian countries and the U.S.(AIO countries). The original sectoral disaggregation covers 76 sectors.

Countries included in the 2000 AIO tables are: China, Indonesia, Japan, Korea, Malaysia, Philippines, Singapore, Taiwan, Thailand and U.S.. The ROW consists of all the 216 countries in the GTAP database (version 6) excluding the AIO countries.

The AIO tables explicitly cover about 56% of world trade excluding trade between countries belonging to the ROW[36]. Bilateral trade flows are especially detailed between and for the AIO countries. Imports from the ROW are available by use for each of the sampled countries. Export to the ROW are not available by use. Countries' reported bilateral flows do not always match with each other. This causes a world deficit of 0.3% the world imports, even after adjusting for the statistical discrepancy[37].

ROW's detailed input-output structures are reconstructed combining the AIO trade data with information in the GTAP database. As a result the 76 AIO sectors and the 57 GTAP ones are aggregated into 34 common sectors. Table B1 describes the 34 sectors in the dataset and the concordance with AIO and GTAP classifications.

AIO countries' GDPs are obtained as the sum of sectoral value added as reported in the input-output tables. The calculation of ROW's income integrates information on the world's income from the World Development Indicators (WDI). Measures of trade barriers are taken from the CEPII dataset. A country's remoteness is calculated as the GDP-weighted sum of that country's distance from any other one in the world.

[36]Trade among AIO countries amounts to 16.63% of world trade. 19.05% (20.16%) of world trade is, instead, from (to) the ROW to (from) the AIO countries. These calculations are based on the GTAP trade data for the year 2001.

[37]All trade data are expressed in f.o.b.. The results in the paper are robust when the trade matrix is constructed to ensure zero world trade balance.

Input-Output Matrix

According to the theory developed in section 1.2, the input-output matrix detailed at bilateral level is defined as follows: $B \equiv \begin{pmatrix} B_{11} & B_{12} & ... & B_{1N} \\ B_{21} & B_{22} & ... & B_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ B_{N1} & B_{N2} & ... & B_{NN} \end{pmatrix}$, where the generic element $B_{ij}(g,h)$ is the amount of good g used for production of one unit of good h, where g is produced by country i and h by country j. For the AIO countries, data on usages of imported intermediates and total output by sector are obtained directly from the AIO tables. ROW's input-output structures are constructed additionally using GTAP data.

Specifically, the domestic input-output structure for the ROW, is derived in two steps. First, I sum the regional input-output tables of all the countries belonging to the ROW to measure the aggregate input-output table for the ROW, $\sum_{j \in ROW} \sum_i M_{ij}(g,h)$. Second, from this table I net out the usage of intermediates imported from the AIO countries and divide the result by the destination sector gross output.

More formally, the matrix of domestic unit intermediates requirements for the ROW, $B_{ROW,ROW}(g,h)$, is calculated as follows:

$$B_{ROW,ROW}(g,h) = \frac{1}{GO(h)} \Big(\sum_{j \in ROW} \sum_i M_{ij}(g,h) - \Big(\frac{M^{IDE}_{AIO,ROW}(g)}{M^{GTAP}_{ROW}(g)} \Big) * \sum_{j \in ROW} \sum_{i \neq j} M_{ij}(g,h) \Big)$$

where $\sum_{j \in ROW} \sum_{i \neq j} M_{ij}(g,h)$ are ROW's imports of good g used in h; $\Big(\frac{M^{IDE}_{AIO,ROW}(g)}{M^{GTAP}_{ROW}(g)} \Big)$ is the share of good g imports from the AIO countries in total ROW's imports of that good[38] with the superscript IDE and $GTAP$ identifying the source of data. ROW's import structures are derived similarly.

Factor Endowments and Factor Intensities

Capital Stocks The GTAP database provides information on the physical capital stock and the accumulation of capital for each region in the database for 2001. These data are adjusted to their 2000 values integrating information from the GTAP database, the AIO

[38] This fraction is not one as each of the GTAP regions composing the ROW imports from every other one. To maximize the consistency of the final dataset, the value of the ROW's imports from the AIO countries is gathered from the AIO tables. The empirical results of the paper are robust if the ROW's imports from the AIO countries is gathered from the GTAP database.

tables and the Penn World Tables (PWT) 6.2.

The standard macro equation for the accumulation of capital implies the value of the capital stock in 2000, K_{2000}, can be written as: $K_{2000} = \frac{(K_{2001}-I_{2000})}{(1-\delta)}$, with I_{2000} and δ being the investment in 2000 and the capital depreciation rate, respectively. The value for the capital stock, K_{2001}, is taken from GTAP, and it is adjusted to account for the change in the price of investment, PI, during 2000-2001. The depreciation rate is set at 13.3% in accordance with the literature (Leamer, 1984). Thus, the capital stock in 2000 is calculated according to: $K_{2000} = \left[K_{2001}^{GTAP} * \frac{PI_{2000}^{PWT6.2}}{PI_{2001}^{PWT6.2}} - I_{2000}\right] * \frac{1}{(1-\delta)}$.

For the AIO countries: gross domestic capital formation data are taken from the AIO tables; the investment price indexes are, instead, obtained from the PWT.

For the ROW the investment series are based on GTAP data for 2001. To retrieve the investment series in 2000 I adjust the GTAP series accounting for the change in investment occurred between 2000 and 2001 using data from the PWT and proceeding in three steps. First, I use PWT data to compute the investment at constant 2000 international prices for years 2000 and 2001 as follows: $I_t^i = RGDPL_t^i * KI_t^i * Pop_t^i$, where $RGDPL^i$ is country i's real per capita GDP using Laspeyres index, KI^i is i's share of real gross domestic investment in $RGDPL^i$ and Pop^i is country i's population. Second, I calculate the change in investment occurred between 2000 and 2001 as the product of the change in real investment occurred between 2000 and 2001, $\frac{I_{2000}^i}{I_{2001}^i}$, and the change in the price of investment over the same period. Finally, for each region in the ROW I measure the investment in 2000 as follows: $\widehat{I_{2000}^i} = I_{2001}^{i,GTAP} * \frac{I_{2000}^{i,PWT6.2}}{I_{2001}^{i,PWT6.2}} * \frac{PI_{2000}^{PWT6.2}}{PI_{2001}^{PWT6.2}}$.

The PWT cover 178 (173) of the 216 countries in the ROW in 2000 (2001). The price index for a GTAP region is the investment-weighted average of investment prices in the countries belonging to that region for which PWT data are available.

Capital Intensities Sectoral payments to capital are equal to the sector value added net of wages and salaries. Following Reimer (2006), the average discount rate across all sectors for a given country i is defined as the ratio of total payments to capital and the country's capital stock: $r_i = \sum_g \frac{\text{payments to capital}_i(g)}{K_i}$. Then, the capital use by sector is simply derived taking the ratio of the sector expenditure on capital and the average discount rate: $b_i^K = \sum_g \frac{\text{payments to capital}_i(g)}{r_i}$. ROW's capital intensities are based on the

GTAP data on firms' payments to capital. Payments are assumed to be stable between 2000 and 2001 up to the investment inflation factor from the PWT.

Labor Endowments Labor endowments for the AIO countries are derived from the employment matrices compiled by the IDE for the year 2000. I use UNPop data on economically active population and the total world population older than 15 to measure the world labor endowment. The labor endowment for the ROW is obtained by difference.

Labor Intensities Labor intensities by sector for the AIO countries are taken from the corresponding employment matrices. The employment data for the ROW are, instead, derived using a method similar to the one followed by Reimer (2006)[39]. Data from the Occupational Wages around the World (OWW) are used to find the employment distribution of a country representative of the ROW. This distribution is then applied to the ROW's total employment to obtain sectoral labor intensities.

The OWW dataset reports data on occupational wages for 161 occupations, 49 industries and 159 countries in 2000. The information on occupational wages is, in many cases, incomplete. Hence, following Reimer, I use the occupational wages of a country representative of the "typical" ROW's exporter to the AIO countries. The "typical" country has to have: i) a per capita income close to that of the ROW's representative exporter to the AIO countries, with the representative exporter's income being calculated as an export-weighted average; ii) a fairly complete occupational information on wages. The per capita GDP of the representative exporter is $16,146 and the selected country is Italy with a per capita GDP of $18,930 in 2000 and wage information for 145 occupations[40].

The Italian employment distribution, in 2000, is obtained taking the ratio between each sectoral wage bill, as obtained from GTAP, and its wage rate. Firms' payments to labor are deflated and adjusted for exchange rate changes, so that data are comparable in 2000 values. Exchange rates data are from the International Financial Statistics (IFS, IMF), and inflation rates from the WDI. The Italian employment distribution is then applied to the total ROW's employment.

By construction, a country's factor content of production equals its factor endowments.

[39] The employment distribution by sector follows similar patterns under the two methods.

[40] The concordance between OWW industries and the 34 sectors in the dataset is available upon request together with further technical details.

Table B1: Sectoral Disaggregation

Description	**Sec. Code**	**AIO code**	**GSC2 No.** (GTAP)
Paddy	1	1	1
Other Grain	2	2	2,3
Cassava, sugar cane and beet, oil palm and coconut,fiber crops other commercial crops, other food crops, oil and fats, sugar	3	3,4 15	4-8, 21,24
Livestock	6	5	9-12
Forestry	7	6	13
Fishery	8	7	14
Crude petroleum and natural gas	9	8	16,17
Copper ore mining, tin ore mining, other non-metallic mineral mining, iron ore, other metaric ore	10	9,10 11	15,18
Slaughtering and meat products	11	4	19,20,22
Rice milling, other grain and flour products, fish and other food products	13	12,13	25,23
Beverage, tobacco	16	16,17	26
Spinning, weaving and dyeing, other made-up textile goods, knitting	17	18-20,22	27
Wearing apparel	18	21	28
Leather and fur, and their products	19	23	29
Lumber, other wood products	20	26,24	30
Pulp, printing and publishing	21	27,28	31
Petroleum and its products	22	34	32
Chemical fertilizers and pesticides, basic industrial chemicals, drugs and medicines, other chemical,products, rubber and plastic products	23	29,30,31-33, 35-37	33
Cement, other non-metallic mineral, glass and glass products	24	38,39,40	34
Basic iron and steel, secondary iron and steel products	25	41	35
Non-ferrous metal	26	42	36
Other metal products	27	43	37
Motor vehicles, air craft, ship building, motor cycles and bicycle, other transportation equipment	28	55,56, 57,58	38,39
Heavy electrical machinery, electronics and its products, other electrical machinery and appliance	30	48,49-51 52,53,54	40
Special industrial machinery, general industrial machinery, engine and turbines, precision instruments	31	44,45 46,47	41
Wooden furniture, other manufacturing products	32	60,25	42
Electricity, gas and water supply	33	61,62	43-45
Building construction, other construction	34	44,45,46,47	46
Wholesale and retail trade	35	65	47
Transportation	36	66	48-50
Telephone and telecommunication	37	67	51
Other business services	38	68	52,53
Other services, unclassified	39	69,71-74,76	54,55,57
Education, Research and Public administration	40	70,75	56